

# Generalizing the Markov and covariance interpolation problem using input-to-state filters

Per Enqvist

**Abstract**—In the Markov and covariance interpolation problem a transfer function  $W$  is sought that match the first coefficients in the expansion of  $W$  around zero and the first coefficients of the Laurent expansion of the corresponding spectral density  $WW^*$ . Here we solve an interpolation problem where the matched parameters are the coefficients of expansions of  $W$  and  $WW^*$  around various points in the disc. The solution is derived using input-to-state filters and is determined by simple calculations such as solving Lyapunov equations and generalized eigenvalue problems.

## I. INTRODUCTION

The problem of designing filters from covariances and Markov parameters has been studied before in numerous papers [17], [13], [14], [15], [16], [20], [21], [18]. Skelton *et. al.* call a stable model matching Markov parameters  $H_0, H_1, \dots, H_{q-1}$  and covariances  $R_0, R_1, \dots, R_{q-1}$  a q-Markov COVariance Equivalent Realization (q-Markov COVER) and they have shown that if the data satisfies a particular consistency condition (which can be avoided using a variable input variance as in [4]), there are many such q-Markov COVERS and they are parameterized by a set of unitary matrices. One of the parameters considered “as known” in the classical q-Markov COVER theory is the variance of the input noise. In [4], [5] the author proposed a method for designing minimal degree realizations using the variance of the input noise as a design parameter which enabled a realization of lower degree to be determined. In fact, that method guarantees a unique stable solution for generic data. Here, using input-to-state filters, we solve an interpolation problem where the matched parameters are the coefficients of expansions of  $W$  and  $WW^*$  around various points in the disc. We could for example consider matching the constraints

$$W(p_1) = q_1, \dots, W(p_n) = q_n$$

for some points  $p_1, \dots, p_n$  in the unit disc and similarly for  $WW^*$ . Most results in [4] carry over to this more general problem. A formal definition of the problem considered is given in the next section. Another approach to this problem was taken in [19]. The main objective of that paper was to prove existence of a fixed point for the Steiglitz-McBride algorithm and a different kind of normalization was used.

## II. PROBLEM FORMULATION

First the Markov and Covariance interpolation problem as formulated in [4] is described and then input-to-state filters are introduced for treating the generalized problem.

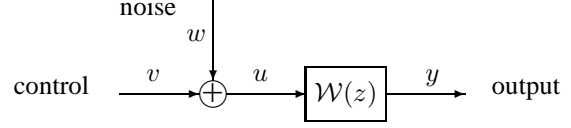


Fig. 1. System considered

### A. The Markov and Covariance interpolation problem

We consider a SISO system where a deterministic control signal  $v$  and a stochastic noise signal  $w$  are fed through the same system  $\mathcal{W}$  to produce the output  $y$  as depicted in Fig. 1.

Define  $u = v + w$  and let  $v$  be the control input and  $w$  an additive noise term. Assuming that the transfer function  $\mathcal{W}$  is rational and of McMillan degree  $\eta$ , it can be described by a minimal state space system

$$\begin{aligned} x_{j+1} &= \mathcal{A}x_j + \mathcal{B}u_j, \\ y_j &= \mathcal{C}x_j + \mathcal{D}u_j, \end{aligned} \quad (1)$$

where  $\mathcal{A} \in \mathbb{C}^{\eta \times \eta}$ ,  $\mathcal{B} \in \mathbb{C}^{\eta \times 1}$ ,  $\mathcal{C} \in \mathbb{C}^{1 \times \eta}$  and  $\mathcal{D} \in \mathbb{C}$ . The output  $y$  is the superposition of the outputs due to each of the inputs  $v$  and  $w$ . Therefore, data from the system can be obtained by the following idealized experiments - or in any other practically more suitable way.

First, determine the output when the noise  $w$  is zero and  $v$  is a unit impulse, yielding the Markov parameters (impulse response parameters)

$$H_0, H_1, \dots, H_\ell. \quad (2)$$

Second, determine the output when the control  $v$  is zero and  $w$  is mean zero white noise with unknown variance  $\Lambda \in \mathbb{C}$ . Assuming  $\mathcal{W}$  is asymptotically stable then this system provides a realization of a stationary stochastic process, and by truncated ergodic sums the covariances

$$R_0, R_1, \dots, R_\ell \quad (3)$$

can be estimated such that the condition

$$\begin{bmatrix} R_0 & R_1 & \ddots & R_\ell \\ \bar{R}_1 & R_0 & \ddots & \ddots \\ \ddots & \ddots & \ddots & R_1 \\ \bar{R}_\ell & \ddots & \bar{R}_1 & R_0 \end{bmatrix} \succeq 0 \quad (4)$$

is satisfied.

### B. Input-to-State filters and interpolation

In order to analyze a signal it is useful to consider a new signal obtained by applying an input-to-state filter [10], i.e. if  $y_k$  is our original signal we define the new state vector  $x_k$  by

$$x_k = Ax_{k-1} + By_k, \quad x_0 = 0, \quad (5)$$

where  $A \in \mathbb{C}^{n \times n}$ ,  $B \in \mathbb{C}^{n \times 1}$  and the eigenvalues of  $A$  lies in the open unit disc. Note that the state  $\chi$  in (1) is not the same as the new “artificial” state  $x$  defined in (5) from the “inputs”  $y_k$ .

Consider the input to state map  $G$ :

$$G(z) := (I - zA)^{-1}B \quad (6)$$

where we will assume that  $(A, B)$  is a reachable pair, i.e.

$$\Gamma = \begin{bmatrix} B & AB & \dots & A^{n-1}B \end{bmatrix} \quad (7)$$

is full rank.

A wide class of interpolation problems can now be approached in a unified framework by expressing the interpolation constraints as inner products with the input-to-state map  $G$ . Let  $\langle \cdot, \cdot \rangle$  denote the standard  $L_2$  inner product on the circle, and for vector- and matrix-valued functions  $F_1$  and  $F_2$  define

$$\langle F_1, F_2 \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} F_1(e^{i\theta}) F_2^*(e^{i\theta}) d\theta,$$

where the integral is evaluated elementwise and  $F^*$  denotes the adjoint of  $F$ , i.e.

$$F^*(z) = F(\bar{z}^{-1})^*,$$

where the superscript  $*$  denotes the usual complex conjugate. Notice that we will allow the  $L_2$  inner product between two matrix-valued functions  $F_1$  and  $F_2$ , possibly of different sizes, provided that the product  $F_1(z)F_2^*(z)$  is well defined.

In the special case where  $A = \text{diag}(p_1, p_2, \dots, p_n)$ ,  $|p_j| < 1$ , and  $B = [1, 1, \dots, 1]^T$ , the scalar function on the  $k$ :th row of  $G(z)$  is

$$g_k(z) = \frac{1}{1 - p_k z}.$$

Then from the Cauchy’s integral formula  $\langle f, g_k \rangle = f(p_k^*)$ , i.e. the values of  $f$  at the selected points can be expressed in terms of the inner product.

In the other special case where all interpolation points are at the origin, i.e. the values of the function and its derivatives at zero are interpolated as in the Caratheodory interpolation problem, then we could chose

$$A = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad (8)$$

so that

$$G(z) = \begin{bmatrix} 1 \\ z \\ \vdots \\ z^{n-1} \end{bmatrix}. \quad (9)$$

The states are then the  $n$  most recent outputs and it is easy to see that the covariance of the state is a Toeplitz matrix as in (4).

In practice one could be interested in having a mixture of interpolation conditions on the function values at different points and on some of its derivatives, and this can be accomplished by considering for example  $A$ -matrices with some particular Jordan structure. To be able to find a  $B$  such that  $(A, B)$  is reachable it is necessary that  $A$  is cyclic, so there can not be more than one Jordan block for each interpolation point (eigenvalue of  $A$ ).

Now given some  $G$ , if  $d\mu$  is a matricial spectral measure of the input (i.e.  $y$  the input to  $G$ ) the state covariance  $\Sigma$  will satisfy [10]

$$\Sigma = \int_{-\pi}^{\pi} G(e^{j\theta}) d\mu(\theta) G(e^{j\theta})^*. \quad (10)$$

For the more general input-to-state filter it is more difficult to know what is the structure of the state-covariance matrix. In Theorem 2 below, a result from [10] describing the feasible structures is stated, but first we need to remind the reader of a well-known result.

*Lemma 1:* The matrix  $\mathcal{G}$  defined by

$$\mathcal{G} \triangleq \langle G, G \rangle, \quad (11)$$

is the Reachability Gramian solving the discrete time Lyapunov equation

$$\mathcal{G} = A\mathcal{G}A^* + BB^*. \quad (12)$$

Since  $(A, B)$  is assumed to be a reachable pair,  $\mathcal{G}$  is invertible.

*Proof:* Note first that

$$AG(z) = z^{-1}(G(z) - B), \quad (13)$$

and then multiply (11) with  $A$  from the left and  $A^*$  from the right to obtain

$$\begin{aligned} A\mathcal{G}A^* &= \langle AGG^*A^*, 1 \rangle \\ &= \langle z^{-1}(G - B)z(G^* - B^*), 1 \rangle \\ &= \langle GG^* + BB^* - BG^* - GB^*, 1 \rangle \\ &= \mathcal{G} + BB^* - BB^* - BB^*. \end{aligned} \quad (14)$$

The last step follows by observing that  $G$  is analytic in the unit disc and thus  $\langle G, 1 \rangle = G(0) = B$ , and similarly  $\langle G^*, 1 \rangle = G^*(\infty) = B^*$ .

Since  $A$  is assumed to be asymptotically stable the solution to the Lyapunov equation is unique, and this completes the proof. ■

*Theorem 2:* A positive definite matrix  $\Sigma$  is a state-covariance matrix for a suitable input process if and only if it is of the form

$$\Sigma = \frac{1}{2}(\mathcal{M}\mathcal{G} + \mathcal{G}\mathcal{M}^*)$$

for a matrix  $\mathcal{M}$  which commutes with  $A$ . Furthermore, any such matrix  $\mathcal{M}$  is uniquely defined modulo an additive imaginary constant  $\alpha I$  with  $\alpha \in j\mathbb{R}$ .

Another way to describe the structure of the state covariance  $\Sigma$  is that it satisfies the equation [8]

$$\Sigma - A\Sigma A^* = BL + L^*B^*$$

for some  $L$ .

Let  $\mathcal{H}_2$  denote the Hardy space of functions that are analytic in the unit disc with square-integrable radial limits, and define

$$\mathcal{K} \triangleq \mathcal{H}_2 \ominus b(z)\mathcal{H}_2, \quad (15)$$

where  $b(z) = \det(zI - A^*)/\det(I - zA)$  is a Blaschke product with poles at the eigenvalues of  $A$ . In fact,  $b(z)$  is the inner, or Douglas-Shapiro-Shields, factor of  $G(z)$ . Then  $\mathcal{K}$  contains all functions in  $\mathcal{H}_2$  which are orthogonal to those that vanish on the spectrum of  $A^*$ , and it is usually called the coinvariant subspace. By [9, Prop. 4] the elements of  $G(z)$  form a basis for  $\mathcal{K}$ , so any  $f \in \mathcal{K}$  can be written  $f(z) = CG(z)$  for some vector  $C$ , and then

$$f(z) = \frac{\det(I - z(A - BC)) - \det(I - zA)}{\det(I - zA)} \in \mathcal{K}. \quad (16)$$

We also need to take inner products between elements in  $\mathcal{K}$  and  $\mathcal{H}_2$ , and then the following formulas are useful.

*Lemma 3:* If  $f(z) \in \mathcal{H}_2$  then

$$\langle f, G \rangle = B^*f(A^*) \quad (17)$$

and

$$\langle G, f \rangle = \bar{f}(A)B. \quad (18)$$

Furthermore,

$$\langle G, fG \rangle = \bar{f}(A)G. \quad (19)$$

Note that it is important here that  $f$  is a scalar function. *Proof:* Since  $f \in \mathcal{H}_2$  and  $G \in \mathcal{H}_2^{n \times 1}$  they have series expansions

$$f(z) = \sum_{k=0}^{\infty} f_k z^k,$$

and

$$G(z) = \sum_{k=0}^{\infty} G_k z^k = \sum_{k=0}^{\infty} A^k B z^k.$$

Then

$$\langle G, f \rangle = \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \langle A^k B z^k, f_{\ell} z^{\ell} \rangle = \sum_{k=0}^{\infty} \bar{f}_k A^k B = \bar{f}(A)B,$$

and the formula for  $\langle f, G \rangle$  follows by considering the complex conjugate.

Finally,

$$\begin{aligned} \langle G, fG \rangle &= \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \sum_{m=0}^{\infty} \langle A^k B z^k, f_{\ell} z^{\ell} G_m z^m \rangle \\ &= \sum_{\ell=0}^{\infty} \sum_{m=0}^{\infty} A^{\ell+m} B \bar{f}_{\ell} G_m^* \\ &= \sum_{\ell=0}^{\infty} \bar{f}_{\ell} A^{\ell} \sum_{m=0}^{\infty} A^m B B^* (A^m)^* \\ &= \bar{f}(A)G, \end{aligned}$$

which concludes the proof.  $\blacksquare$

*Remark 4:* Alternatively, this could be proven by considering generalized Cauchy kernels

$$\langle G, f \rangle = \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} \bar{f}(e^{-j\theta})(I - e^{j\theta}A)^{-1} d\theta \right) B,$$

as in [10].

Estimation of the parameters from data can be performed by applying the input-to-state filter and then using standard techniques, see [2] for examples of filter bank data analysis.

### III. THE GLOBAL OPTIMIZATION PROBLEM

We will assume here that the spectral measure in (10) is given by

$$d\mu(\theta) = W(e^{i\theta}) \Lambda d\theta W(e^{i\theta})^*,$$

where

$$W(z) = \sum_{k=0}^{\infty} w_k z^k \in \mathcal{H}_2, \quad (20)$$

i.e. is analytic in the unit disc (so the sum converges for all  $z$  in the unit disc), and in this class of spectral measures we will find the one allowing the maximal input variance  $\Lambda$  meanwhile satisfying the following interpolation conditions:

$$\langle G W \Lambda W^* G^*, 1 \rangle = \Sigma, \quad (21)$$

where the state covariance  $\Sigma$  satisfies the condition in Theorem 2, and

$$\langle G, W \rangle = H, \quad (22)$$

for an arbitrary nonzero state-Markov vector  $H$ .

The interpolation constraint in (21) was considered in, for example, [8]. The interpolation constraint in (22) can be recognized as a special case of the Lagrange-Sylvester interpolation as studied in [1, section 16]. Here, both constraints are enforced simultaneously.

Thus the optimization problem considered is:

$$(\mathfrak{S}) \quad \left[ \begin{array}{l} \max \\ W \in \mathcal{H}_2 \\ \Lambda \in \mathbb{R}^+ \\ \text{s.t.} \quad \left\{ \begin{array}{l} \langle G W \Lambda W^* G^*, 1 \rangle = \Sigma, \\ \langle G, W \rangle = H. \end{array} \right. \end{array} \right]$$

Let  $\Xi$  be an  $(n \times n)$  Hermitian matrix and  $\zeta$  be an  $(1 \times n)$  vector consisting of Lagrange multipliers, the Lagrangian is then

$$\begin{aligned} \mathcal{L}(W, \Lambda) &\triangleq \Lambda + \text{tr}\{(\Sigma - \langle G W \Lambda W^* G^*, 1 \rangle)\Xi\} \\ &\quad + \zeta(\langle G, W \rangle - H). \end{aligned}$$

We can rewrite it in the following form

$$\begin{aligned} \mathcal{L}(W, \Lambda) &= \Lambda + \text{tr}\{\Sigma \Xi\} - \langle W \Lambda W^*, G^* \Xi G \rangle \\ &\quad - \zeta H + \langle \zeta G, W \rangle. \end{aligned}$$

where

$$G^* \Xi G = B^*(I - \bar{z}^{-1}A^*)^{-1}\Xi(I - zA)^{-1}B$$

and

$$\zeta G = \zeta(I - zA)^{-1}B$$

are scalar functions.

Before taking the maximum we write it in the form

$$\mathcal{L}(W, \Lambda) = \langle \zeta G, W \rangle + \langle (1 - W^* G^* \Xi G W) \Lambda, 1 \rangle + \text{tr}\{\Sigma \Xi\} - \zeta H$$

**Note:**  $\text{Sup}\{\mathcal{L}(W, \Lambda) | W \in \mathcal{H}_2, \Lambda > 0\} < \infty$  only if  $G^* \Xi G$  is in the “positive cone”, i.e. it is non-negative for all  $z$  on the unit circle, and

$$\langle W^* G^* \Xi G W, 1 \rangle \geq 1. \quad (23)$$

Maximizing over  $\Lambda$  while assuming (23) it must hold that

$$\Lambda \langle 1 - W^* G^* \Xi G W, 1 \rangle = 0, \quad (24)$$

and since  $\Lambda \neq 0$ , equality must hold in (23), i.e.

$$\langle W^* G^* \Xi G W, 1 \rangle = 1. \quad (25)$$

Maximizing over  $W$  shows that the following variation has to be zero for all  $\delta W$

$$\langle \zeta G - 2\Lambda G^* \Xi G W, \delta W \rangle = 0. \quad (26)$$

Therefore

$$\zeta G = 2\Lambda G^* \Xi G W + V^* \quad (27)$$

where  $V \in \mathcal{H}_2$  and  $V(0) = 0$ . From (27) the poles of  $V^*$  has to be poles of  $G^*$ . Furthermore,  $V \in \mathcal{K}$  follows by considering the partial fraction expansion of (27), so there must be a vector  $\nu$  such that

$$V(z) = \nu G(z) \quad (28)$$

Then the transfer function  $W$  will be given by

$$\begin{aligned} W &= \frac{1}{2} (G^* \Xi G)^{-1} (\zeta G - V^*) \Lambda^{-1} \\ &= \frac{1}{2\Lambda} (G^* \Xi G)^{-1} (\zeta G - G^* \nu^*). \end{aligned}$$

*Lemma 5:* If  $\Xi$  is non-negative we can factor  $G^* \Xi G$  as

$$G^* \Xi G = (\xi G)^* (\xi G), \quad (29)$$

where  $\xi$  is a row-vector.

*Proof:* Since  $\Xi$  is non-negative and Hermitian it can be factorized as

$$\Xi = \begin{bmatrix} \xi_1^* & \xi_2^* & \cdots & \xi_n^* \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix}.$$

Then

$$G^* \Xi G = \sum_{k=1}^n G^* \xi_k^* \xi_k G \quad (30)$$

and  $G^* \Xi G$  is a sum of elements in  $\mathcal{K} \cup \mathcal{K}^*$ , where  $k^* \in \mathcal{K}^*$  if  $k \in \mathcal{K}$ . Since all the terms in (30) are positive, by spectral factorization a vector  $\xi$  such that the sum is equal to  $G^* \xi^* \xi G$  can be found. ■

For  $W$  to be analytic outside the unit disc it is necessary that the factor  $(\xi G)^{-*}$  is cancelled, i.e. we need that

$$W = \frac{1}{2\Lambda} \frac{\zeta G - V^*}{(\xi G)(\xi G)^*} = \frac{\sigma G}{\xi G}.$$

From (27) and (29) it follows that

$$\zeta G = 2\Lambda G^* \xi^* \sigma G + G^* \nu^*. \quad (31)$$

Using (25) and (31) the dual function is

$$\begin{aligned} \varphi(\xi, \sigma) &= \left\langle \zeta G, \frac{\sigma G}{\xi G} \right\rangle + \text{tr}\{\Sigma \xi^* \xi\} - \zeta H \\ &= 2\Lambda \sigma \mathcal{G} \sigma^* + \xi \Sigma \xi^* - \zeta H \end{aligned}$$

since  $\langle V^*, W \rangle = 0$  and where  $\mathcal{G}$  was defined in (11).

To determine the last term  $\zeta H$ , multiply (31) with  $G^*$  and integrate to obtain:

$$\zeta \langle G G^*, 1 \rangle = 2\Lambda \langle G^* \xi^* \sigma G G^*, 1 \rangle + \langle V^* G^*, 1 \rangle$$

the last term is zero and then

$$\zeta H = 2\Lambda \sigma \langle G(G^* \mathcal{G}^{-1} H) G^*, 1 \rangle \xi^* = 2\Lambda \sigma \mathcal{H} \xi^*,$$

where

$$\mathcal{H} \triangleq \langle G(G^* \mathcal{G}^{-1} H) G^*, 1 \rangle. \quad (32)$$

*Lemma 6:* The matrix  $\mathcal{H}$  defined by (32) is the unique solution to the Stein equation

$$\mathcal{H} = A \mathcal{H} A^* + H B^*. \quad (33)$$

*Proof:* As in the proof of Lemma 1, note that (13) holds and then multiply (32) with  $A$  from the left and  $A^*$  from the right to obtain

$$\begin{aligned} A \mathcal{H} A^* &= \langle A G G^* A^* (G^* \mathcal{G}^{-1} H), 1 \rangle \\ &= \langle z^{-1} (G - B) z (G^* - B^*) (G^* \mathcal{G}^{-1} H), 1 \rangle \\ &= \langle (G G^* + B B^* - B G^* - G B^*) (G^* \mathcal{G}^{-1} H), 1 \rangle \\ &= \mathcal{H} + B \langle (B^* - G^*) (G^* \mathcal{G}^{-1} H), 1 \rangle \\ &\quad - \langle G B^* (G^* \mathcal{G}^{-1} H), 1 \rangle \end{aligned} \quad (34)$$

The second term in (34) is zero since the integrand is analytic outside the unit circle and  $G^*(\infty) = B^*$ .

The third term in (34) is  $H B^*$ , which follows by considering the action on an arbitrary vector  $v$ ;

$$\begin{aligned} \langle G B^* (G^* \mathcal{G}^{-1} H), 1 \rangle v &= \langle G B^* v (G^* \mathcal{G}^{-1} H), 1 \rangle \\ &= \langle G G^* \mathcal{G}^{-1} H, 1 \rangle B^* v \\ &= H B^* v. \end{aligned}$$

Since  $A$  is assumed to be asymptotically stable the solution to the Stein equation is unique, and this completes the proof. ■

*Remark 7:* Note that if

$$h(z) := (I - zA)^{-1} H \quad (35)$$

then  $\langle h, G \rangle$  solves the Stein equation (33), and then by uniqueness (compare [9, Eq. (40)])

$$\mathcal{H} = \langle h, G \rangle.$$

Now, the dual optimality function is

$$\begin{aligned}\varphi(\xi, \sigma) &= 2\Lambda\sigma\mathcal{G}\sigma^* + \xi\Sigma\xi^* - 2\Lambda\sigma\mathcal{H}\xi^* \\ &= \begin{bmatrix} \sigma & \xi \end{bmatrix} \begin{bmatrix} 2\Lambda\mathcal{G} & -2\Lambda\mathcal{H} \\ 0 & \Sigma \end{bmatrix} \begin{bmatrix} \sigma^* \\ \xi^* \end{bmatrix} \quad (36)\end{aligned}$$

Maximizing this expression over positive  $\Lambda$

$$\varphi(\xi, \sigma) = 2\Lambda(\sigma\mathcal{G}\sigma^* - \sigma\mathcal{H}\xi^*) + \xi\Sigma\xi^*$$

it is clear that  $(\sigma\mathcal{G}\sigma^* - \sigma\mathcal{H}\xi^*)$  has to be zero. In fact, if it is negative the optimal value of  $\Lambda$  would be zero and we have assumed that it is positive, and if it is positive the optimal value of  $\Lambda$  would be infinite. Furthermore, the following holds:

*Lemma 8:* Given that  $W(z) = \frac{\sigma G(z)}{\xi G(z)} \in \mathcal{H}_2$ , the constraint  $\langle G, W \rangle = H$  implies that  $\mathcal{H}\xi^* = \mathcal{G}\sigma^*$ .

*Proof:* We know that  $\langle G, W \rangle = H$ , and thus

$$G^*\mathcal{G}^{-1} \left\langle G, \frac{\sigma G}{\xi G} \right\rangle = G^*\mathcal{G}^{-1}H,$$

is a scalar function, so

$$\begin{aligned}\mathcal{H}\xi^* &= \langle GG^*(G^*\mathcal{G}^{-1}H), 1 \rangle \xi^* \\ &= \left\langle GG^*\xi^* \left( G^*\mathcal{G}^{-1} \left\langle G, \frac{\sigma G}{\xi G} \right\rangle \right), 1 \right\rangle \\ &= \langle GG^*, \xi G \rangle \mathcal{G}^{-1} \left\langle G, \frac{\sigma G}{\xi G} \right\rangle \\ &= \langle G, (\xi G)G \rangle \mathcal{G}^{-1} \left\langle G, \frac{1}{\xi G} G \right\rangle \sigma^* \\ &= \overline{(\xi G(A))} \mathcal{G} \mathcal{G}^{-1} \overline{(\xi G(A))}^{-1} \mathcal{G} \sigma^* \\ &= \mathcal{G} \sigma^*\end{aligned}$$

where we have used (19) twice. ■

The complementarity condition (25) can be formulated as

$$\left\langle \frac{G^*\sigma^*}{G^*\xi^*} G^* \Xi G \frac{\sigma G}{\xi G}, 1 \right\rangle = \sigma\mathcal{G}\sigma^* = 1,$$

and then the dual problem is

$$(\mathfrak{D}) \quad \begin{bmatrix} \min_{\sigma, \xi} & \xi\Sigma\xi^* \\ \text{s.t.} & \begin{cases} \mathcal{G}\sigma^* = \mathcal{H}\xi^*, \\ \sigma\mathcal{G}\sigma^* = 1 \end{cases} \end{bmatrix}$$

where  $\sigma$  and  $\xi$  are related by the Markov interpolation conditions.

The variable  $\Lambda$  was eliminated above, but it can be recovered by considering the dual of the dual. Let  $\Lambda$  be the Lagrange multiplier and use  $\mathcal{G}\sigma^* = \mathcal{H}\xi^*$  to eliminate  $\sigma$

$$\begin{aligned}L &= \xi\Sigma\xi^* - \Lambda(\sigma\mathcal{G}\sigma^* - 1) \\ &= \xi\Sigma\xi^* - \Lambda\xi\mathcal{H}^*\mathcal{G}^{-1}\mathcal{H}\xi^* + \Lambda \\ &= \xi(\Sigma - \Lambda\mathcal{H}^*\mathcal{G}^{-1}\mathcal{H})\xi^* + \Lambda\end{aligned}$$

which leads us to maximize  $\Lambda$  as  $\Sigma - \Lambda\mathcal{H}\mathcal{G}^{-1}\mathcal{H}^*$  is non-negative, i.e.

$$(\mathfrak{P}^*) \quad \begin{bmatrix} \max_{\Lambda} & \Lambda, \\ \text{s.t.} & \Sigma - \Lambda\mathcal{H}^*\mathcal{G}^{-1}\mathcal{H} \succeq 0 \end{bmatrix}$$

The optimal  $\Lambda$  is now given by the largest positive value such that  $\Sigma - \Lambda\mathcal{H}^*\mathcal{G}^{-1}\mathcal{H}$  is non-negative definite, i.e. the smallest generalized eigenvalue of  $(\Sigma, \mathcal{H}^*\mathcal{G}^{-1}\mathcal{H})$ .

*Theorem 9:* Given a state covariance  $\Sigma$  satisfying the condition in Theorem 2, and an arbitrary nonzero state-Markov vector  $H$ . Then, an optimizer  $W$  to problem (S) is given by  $W(z) = (\sigma G(z))/(\xi G(z))$ , where  $\xi$  is a nonzero solution to the equation

$$(\Sigma - \Lambda\mathcal{H}^*\mathcal{G}^{-1}\mathcal{H})\xi^* = 0,$$

$\Lambda$  is the smallest generalized eigenvalue of  $(\Sigma, \mathcal{H}^*\mathcal{G}^{-1}\mathcal{H})$ , and finally  $\sigma$  is determined by

$$\sigma^* = \mathcal{G}^{-1}\mathcal{H}\xi^*.$$

Furthermore, if the smallest generalized eigenvalue has multiplicity one the optimizer  $W$  is unique.

*Proof:* This follows from the derivation above. ■

As in the Markov and Covariance interpolation problem there is a special choice of the Markov parameters that reduce the problem to the equivalent of the Pisarenko method [11]. Namely, choosing  $H = B$ , then  $\mathcal{H} = \mathcal{G}$  and maximizing  $\Lambda$  under the constraint

$$\Sigma - \Lambda\mathcal{H}^*\mathcal{G}^{-1}\mathcal{H} = \Sigma - \Lambda\mathcal{G} \succeq 0$$

makes the corresponding  $W$  an inner function [11].

#### IV. MODEL REDUCTION EXAMPLE

To illustrate the method proposed here, a model reduction application is considered. The method proposed here is a generalization of the q-Markov COVER methods, that were initially proposed to be used for model reduction [14].

The transfer function from input 2 to output 1 of a portable CD-player is considered. This model, of order 120, is provided by SLICOT [3], and has been used by, for example, [12], [7]. The magnitude of the transfer function is depicted with a thick solid green line in Figure 2. There is a wide range of frequencies over which there are interesting features of the Bode plot.

The given transfer function is a continuous time stable function. The bilinear map

$$z = \frac{1 - sT/2}{1 + sT/2}, \quad (37)$$

where  $T = 1/250$ , is used to transform the continuous time model into a discrete time model.

Using three different input-to-state filters, reduced order models of degree 13 are designed. Our aim will not be to find the optimal interpolation point locations for this particular model, but to illustrate the way this choice effects the solutions.

First an input-to-state filter as in (8) was applied, corresponding to the Markov and covariance interpolation problem described in section II-A, and the magnitude plot of the resulting model is depicted with a blue dashed line in Figure 2.

Then, an input-to-state filter with 14 poles spread evenly around a circle with radius 0.95 was applied. The magnitude

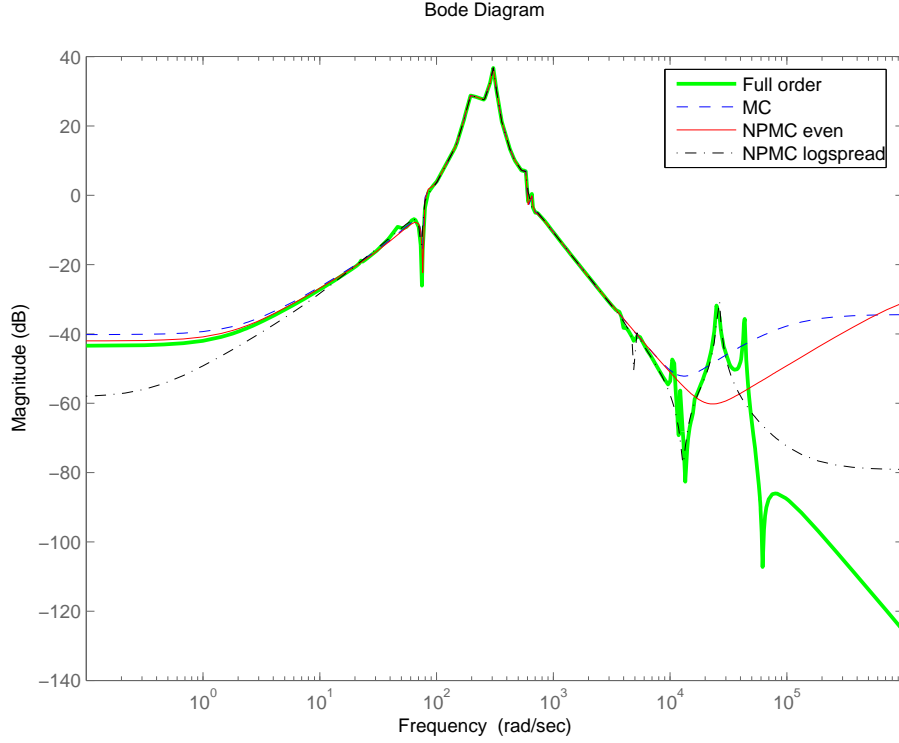


Fig. 2. Model reduction

plot, depicted in Figure 2 with a solid red line, is similar to the first one, but with a slightly smaller error for low and high frequencies.

Finally, an input-to-state filter with 14 poles spread unevenly around a circle with radius 0.9 was applied. The spread in frequencies were chosen to correspond to a logarithmic spread in the frequency interval  $10^1$  to  $10^5$ . In the discrete domain, the interpolation point locations are depicted with black plusses in Figure 3 together with the interpolation points of the two other filters. This choice of interpolation points is made to compensate for the frequency warping caused by the bilinear map (37). The magnitude plot, depicted in Figure 2 with a black dashed-dotted line, shows an improvement of the fit in the frequency range where the poles were chosen.

For comparison, a model of degree 13 is determined using a standard balanced truncation model reduction method and its magnitude plot is depicted in Figure 4. A good fit for the interval of frequencies where the magnitude is large is obtained. It is well known that weights can be applied to improve the fit for certain frequency regions. The choice of these weights, as well as the choice of interpolation points in our approach, should be made with the prior knowledge and requirements of the low order model in mind.

## V. USEFUL FORMULAS FOR THE USER

In this section we give simplified formulas for calculating the transfer function and state space representations of the  $W$  parameterized by  $\xi$  and  $\zeta$ . It is also shown how to determine the state-covariances and state-Markov parameters

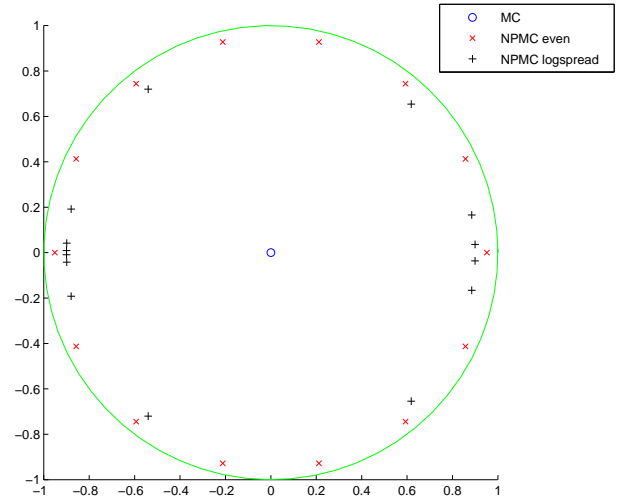


Fig. 3. Interpolation point locations

from these representations of  $W$ . (Note that the problem considered in this paper is the inverse of determining the interpolation parameters from the model  $W$ .) These formulas will be important for applying the method proposed here.

### A. For transfer functions

Since  $W = \sigma G / \xi G$  is a quotient between two functions in  $\mathcal{K}$ , it follows from (16) that it can be written as a quotient of two polynomials

$$W(z) = \frac{b(z)}{a(z)} \quad (38)$$

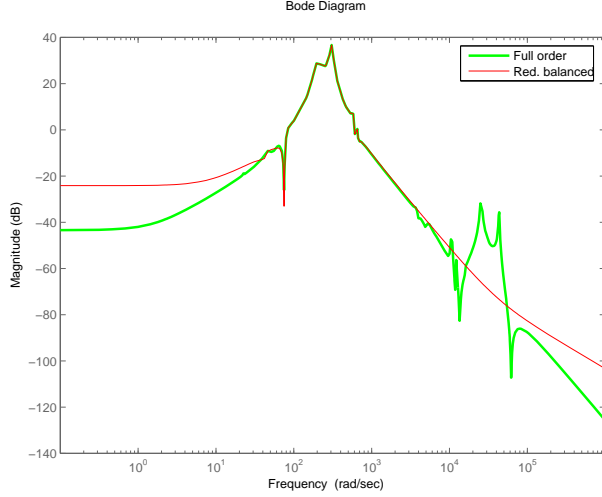


Fig. 4. Model reduction

where

$$b(z) = \det(I - z(A - B\sigma^*)) - \det(I - zA) \quad (39)$$

and

$$a(z) = \det(I - z(A - B\xi^*)) - \det(I - zA). \quad (40)$$

Clearly,  $z = 0$  is a zero of both  $a$  and  $b$  so it is cancelled out, which leaves a  $W$  of degree at most  $n - 1$ .

To check if the resulting transfer function  $W$  satisfies the interpolation conditions it is convenient to use (18) to obtain

$$\langle G, W \rangle = \bar{W}(A)B = \bar{b}(A)\bar{a}(A)^{-1}B.$$

To determine the state-covariance  $\mathcal{S}$  corresponding to a particular  $W$  we can use the following formula from [10]

$$\mathcal{S} = \Lambda \langle GW, GW \rangle = \frac{1}{2} \Lambda (\Psi \mathcal{G} + \mathcal{G} \Psi^*) \quad (41)$$

where  $\Psi = \bar{f}(A)$  and  $f$  is the positive real part of  $WW^*$ . The function  $f$  satisfies

$$f + f^* = WW^*,$$

and it is clear that  $f = d/a$ , where  $a$  is given by (40) and  $d$  solves the equation

$$b(z)b^*(z) = d(z)a^*(z) + a(z)d^*(z). \quad (42)$$

This equation has a unique solution  $d$  such that all the roots of  $d$  are outside the unit disc provided that all the roots of  $a$  also are outside the unit disc [22], [23].

#### B. For state-space realizations

A state-space realization of degree  $n$  corresponding to (38) is given by

$$W(z) = [z\sigma\beta(I - A\beta z)^{-1}AB + \sigma B](\xi B)^{-1}, \quad (43)$$

where  $\beta = I - B\xi(\xi B)^{-1}$ . From the last section it is known that a state-space realization of  $W$  of degree  $n - 1$  exists, and one is determined in the appendix.

Given a state space representation of  $\mathcal{W}$  as in (1) the product  $GW$  has a realization

$$\left( \begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array} \right) = \left( \begin{array}{cc|c} \mathcal{A} & 0 & \mathcal{B} \\ \hline BC & A & BD \\ 0 & I & 0 \end{array} \right). \quad (44)$$

Then the state-Markov parameter is given by

$$\langle G, \mathcal{W} \rangle = BD^* + A\tilde{P}C^*, \quad (45)$$

where  $\tilde{P}$  solves the Stein equation  $\tilde{P} = A\tilde{P}A^* + BB^*$ , and the state-covariance  $\mathcal{S}$  is given by

$$\mathcal{S} = \Lambda \hat{C} \hat{P} \hat{C}^*, \quad (46)$$

where  $\hat{P}$  solves the Lyapunov equation  $\hat{P} = \hat{A}\hat{P}\hat{A}^* + \hat{B}\hat{B}^*$ .

## VI. CONCLUSIONS AND FUTURE WORK

The ideas and results in [5] were shown to carry over to the case where not all interpolation points are at zero. This freedom of choosing the interpolation points can be used to obtain an improved matching at some frequency regions. One example was given to illustrate the effect of moving the interpolation points. Input-to-state filters proved to be a convenient tool to derive this theory and simple formulas based on solving Lyapunov equations were obtained. However, if really high order models are considered specialized numerical tools have to be developed.

The approach used in [7] applies only the interpolation on  $\Sigma$  and instead of matching  $H$  arbitrary spectral zeros may be chosen. This gives the user more freedom in designing the model, but at the price of having to tune more parameters.

In [6], a generalization of the Markov and covariance interpolation problem with variable input variance to MIMO systems was considered. A similar generalization should be possible here.

## VII. ACKNOWLEDGEMENT

The author wishes to thank the anonymous referee providing valuable comments.

## APPENDIX

We first derive (43)

$$\begin{aligned} W(z) &= \frac{\sigma(I - zA)^{-1}B}{\xi(I - zA)^{-1}B} \\ &= \frac{\sigma B + \sigma(z^{-1}I - A)^{-1}AB}{\xi B + \xi(z^{-1}I - A)^{-1}AB} \\ &= (\sigma B + \sigma(z^{-1}I - A)^{-1}AB) \times \\ &\quad (I - \xi(z^{-1}I - A\beta)^{-1}AB(\xi B)^{-1})(\xi B)^{-1} \\ &= \left( \begin{array}{cc|c} A\beta & 0 & AB(\xi B)^{-1} \\ \hline -AB\xi & A & AB \\ -\sigma B\xi & \sigma & \sigma B \end{array} \right) (\xi B)^{-1} \\ &= \left( \begin{array}{cc|c} A\beta & 0 & AB(\xi B)^{-1} \\ \hline 0 & A & 0 \\ \sigma B\xi\beta & \sigma & \sigma B \end{array} \right) (\xi B)^{-1} \\ &= \left( \begin{array}{c|c} A\beta & AB \\ \hline \sigma\beta & \sigma B \end{array} \right) (\xi B)^{-1}. \end{aligned}$$

Here we have used the matrix

$$T = \begin{bmatrix} I & 0 \\ -(\xi B)I & I \end{bmatrix}$$

to do a change of basis in the large system before cancelling the unreachable second part of the state vector.

This realization is still non-minimal since there is both a zero and a pole at infinity. Note that

$$\beta B = B - (\xi B)^{-1} B \xi B = 0. \quad (47)$$

Now we use  $\Gamma$  in (7) to do a change of basis. From (47) it follows that

$$\sigma \beta \Gamma = \sigma \beta A \begin{bmatrix} 0 & B & \cdots & A^{n-2} B \end{bmatrix},$$

$$\Gamma^{-1} A \beta \Gamma = \Gamma^{-1} A \beta A \begin{bmatrix} 0 & B & \cdots & A^{n-2} B \end{bmatrix},$$

and

$$\Gamma^{-1} A B = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \end{bmatrix}^T.$$

Then the first state in the new basis is not observable or reachable so a reduced order realization is obtained by cancelling it:

$$W(z) = \left( \frac{\hat{\Gamma} A \beta A \tilde{\Gamma}}{\sigma \beta A \tilde{\Gamma}} \middle| \frac{e_1}{\sigma B} \right) (\xi B)^{-1}, \quad (48)$$

where

$$\tilde{\Gamma} \triangleq \begin{bmatrix} B & AB & \cdots & A^{n-2} B \end{bmatrix},$$

and

$$\hat{\Gamma} \triangleq (\Gamma^{-1})_{2:n} = \begin{bmatrix} 0 & I \end{bmatrix} \Gamma^{-1},$$

and the subindex  $2 : n$  denotes rows 2 to  $n$  of the matrix.

In particular, if the characteristic polynomial  $\chi_A(t)$  is parameterized as

$$\chi_A(t) = t^n + \chi_1 t^{n-1} + \cdots + \chi_{n-1} t + \chi_n,$$

the dynamics matrix in (48) is

$$\begin{aligned} \tilde{A} &\triangleq \begin{bmatrix} 0 & I \end{bmatrix} \Gamma^{-1} A \beta A \tilde{\Gamma} \\ &= \begin{bmatrix} 0 & I \end{bmatrix} \Gamma^{-1} A \left( \tilde{\Gamma} - (\xi B)^{-1} A B \xi \tilde{\Gamma} \right) \\ &= \begin{bmatrix} 0 & -\chi_{n-1} \\ I & -\chi \end{bmatrix} - (\xi B)^{-1} e_1 \xi \tilde{\Gamma} \\ &= \begin{bmatrix} -\gamma & -\gamma_{n-2} - \chi_{n-1} \\ I & -\chi \end{bmatrix} \end{aligned}$$

where

$$\chi \triangleq \begin{bmatrix} \chi_{n-2} & \cdots & \chi_1 \end{bmatrix}^T,$$

$$\gamma \triangleq \begin{bmatrix} \gamma_0 & \cdots & \gamma_{n-3} \end{bmatrix}$$

and  $\gamma_k = \xi A^k B / (\xi B)$  for  $k = 0, 1, \dots, n-2$ .

## REFERENCES

- [1] J.A. Ball, I. Gohberg and L. Rodman, "Interpolation of rational matrix functions". *Operator Theory: Advances and Applications*, 45, Birkhäuser, 1990.
- [2] C. Byrnes, T. Georgiou and A. Lindquist, "A new approach to spectral estimation: A tunable high-resolution spectral estimator". *IEEE Trans. Signal Processing*, vol. 48, No. 11, pp.3189–3205, 2000.
- [3] Y. Chahlaoui and P. Van Dooren, "A collection of benchmark examples for model reduction of linear time invariant dynamical systems", *SLICOT Working Note 2002-2: February 2002*.
- [4] P. Enqvist, "Pole-zero identification based on simultaneous realization of normalized covariances and Markov parameters" in *Proceedings CDC 2002*
- [5] P. Enqvist, "On the simultaneous realization problem: Markov parameter and covariance interpolation". *Signal Processing*, vol. 86, No. 10, pages 3043–3054, Oct. 2006.
- [6] P. Enqvist, "On Minimal MIMO Realization of Covariance and Markov Parameters". *Proc. MTNS, Kyoto, 2006*.
- [7] G. Fanizza, J. Karlsson, A. Lindquist and R. Nagamune, "Passivity-preserving model reduction by analytic interpolation". *Linear Algebra and Applications*, to be published.
- [8] T. Georgiou, "Spectral analysis based on the state covariance: The maximum entropy spectrum and linear fractional parametrization". *IEEE Trans. Automatic Control*, vol. 47, No. 11, pages 1811 – 1823, Nov. 2002.
- [9] T. Georgiou, "The structure of state covariances and its relation to the powerspectrum of the input". *IEEE Trans. Automatic Control*, vol. 47, No. 7, pages 1056 – 1066, Jul. 2002.
- [10] T. Georgiou, "Spectral Estimation via Selective Harmonic Amplification". *IEEE Trans. Automatic Control*, vol. 46, No. 1, pages 29 – 42, Jan. 2001.
- [11] T. Georgiou, "The Carathéodory-Fejér-Pisarenko decomposition and its multivariable counterpart". *IEEE Trans. Automatic Control*, vol. 52, No. 2, Feb. 2007, pages 212 – 228.
- [12] S. Gugercin and A.C. Antoulas, "Model reduction of large scale systems by least squares", *Linear Algebra and its Applications, Special Issue on Order Reduction of Large-scale Systems*, 2004
- [13] Y. Inouye, "Approximation of multivariable linear systems with impulse response and autocorrelation sequences". *Automatica*, 19(3):265–277, 1983.
- [14] A.M. King, U.B. Desai, and R.E. Skelton, "A generalized approach to q-Markov covariance equivalent realizations for discrete systems". *Automatica*, 24(4):507–515, 1988.
- [15] K. Liu and R.E. Skelton, "A new formulation of Q-Markov covariance equivalent realization". *Applied Mathematics and Computation*, 53:83–95, 1993.
- [16] K. Liu and R.E. Skelton, "Q-Markov covariance equivalent realization and its applications to flexible structure identification". *Journal of Guidance, Control and Dynamics*, 16(2):308–319, March–April 1993.
- [17] C.T. Mullis and R.A. Roberts, "The use of second-order information in the approximation of discrete-time linear systems". *IEEE Trans. on Acoustics, Speech and Signal Processing*, ASSP-24(3):226–238, 1976.
- [18] B. Musicus and A. Kabel, "Maximum entropy pole-zero estimation". *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 11, pp: 1389–1392, 1986.
- [19] P.A. Regalia, M Mboup and M Ashari-Astani "A class of first- and second-order interpolation problems in model reduction". *Archiv für Elektronik und Übertragungstechnik*, vol. 49, no. 5/6, pp. 332–343, 1995.
- [20] R.E. Skelton and B.D.O. Anderson, "The generation of all Q-Markov COVERS". *IEEE Circuits and Systems*, 35(4):375–384, 1988.
- [21] R.E. Skelton and G. Shi, "Iterative identification and control using a weighted q-Markov cover with measurement noise". *Signal Processing*, 52:217–234, 1996.
- [22] G. Wilson, "Factorization of the Covariance Generating Function of a Pure Moving Average Process". *SIAM J. Numer. Anal.*, vol 6, no. 1, pp. 1–7, 1969.
- [23] Tim N.T. Goodman, Charles A. Micchelli, Giuseppe Rodriguez and Sebastiano Seatzu "Spectral factorization of Laurent polynomials" *Advances in Computational Mathematics*, 7, pp. 429–455, 1997.